

Contents lists available at [ScienceDirect](http://www.sciencedirect.com)

Contemporary Clinical Trials Communications

journal homepage: www.elsevier.com/locate/conctc

Blinding properties of methods for supplying drug kits to investigational sites

Ron Yu ^{a,*}, Daniel A. Coleman ^b^a Clinical Biostatistics, Genentech, Inc., United States^b Nonclinical Biostatistics, Genentech, Inc., United States

ARTICLE INFO

Article history:

Received 9 July 2015

Received in revised form

3 October 2015

Accepted 13 October 2015

Available online 26 October 2015

Keywords:

Drug kit

Supply method

Partial unblinding

Blinded group ordering

Strong blinding

ABSTRACT

Most blinded, late stage, randomized clinical trials package study drug, active or placebo, into drug kits for distribution to investigational sites. Drug kits enable investigators to administer study drug to subjects in a blinded manner without the assistance of an unblinded pharmacist. Supply methods determine when and how many kits to send to sites. If not properly designed, these methods can partially unblind investigators, i.e., investigators can conclude that two subjects are (1) on the same treatment arm with certainty or (2) on different treatment arms with certainty. Partial unblinding can bias the way investigators provide patient care, report adverse events and assess efficacy endpoints, and can lead to full unblinding when the other subject is unblinded. In this paper, we describe several examples of partial unblinding in the supply methods commonly used by many Randomization and Trial Supply Management (RTSM) systems, propose a new criterion for evaluating the blinding properties of supply methods, and prove that two alternative supply methods do not permit full or partial unblinding, even after the investigator is unblinded up to a certain number of other subjects.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Ensuring sufficient study drug supplies at sites is a requirement of every clinical trial. The primary difficulty lies in the requirement of maintaining a standing inventory of drug kits at sites to dose subjects when subjects are randomized. This requirement stems from the common practice of dosing subjects on the same day or within 1–2 days of randomization in order to prevent early withdrawals as well as provide subject convenience. Under these circumstances, there is not sufficient time to ship a kit to the site on an as-needed basis. Instead a standing inventory of kits must be maintained at sites. The size of this inventory must be large enough to accommodate subject randomizations that are not known ahead of time but not too large to cause a significant waste.

For blinded studies, ensuring the blind is also of crucial importance. There are two kinds of unblinding: full unblinding and partial unblinding [1]. Full unblinding occurs when an investigator learns the kit type of a particular kit. Partial unblinding occurs when an investigator can conclude with certainty that two kits are of the same kit type or conclude with certainty that two kits are of

different kit types. Both full and partial unblinding should be avoided.

To prevent unblinding by investigators, randomness should be incorporated in each step of the “life-cycle” of a drug kit, including packaging and labeling, supplying to sites and assigning to subjects. Drug kit labeling methods and their blinding properties have been discussed extensively in the literature [1–3]. In regard to assigning kits to subjects, it is known that deterministic methods such as assigning kits from the earliest shipments first (i.e., first-in-first-out [FIFO]) can partially unblind investigators [4]. This paper describes several methods for supplying kits and compares their blinding properties.

This paper is organized as follows. Section 2 describes two supply strategies to illustrate the trade-off between ensuring the blind and kit efficiency (i.e., total number of kits required for the trial; supply methods that require fewer kits for the trial are more efficient). Sections 3 and 4 describe the widely used Trigger-and-Resupply (TR) and Predictive Resupply (PR) methods [5,6] and their poor blinding properties. Because Randomization and Trial Supply Management (RTSM) system implementations are proprietary, our descriptions of these methods may be somewhat different from actual vendor implementations. Section 5 proposes a new criterion called *Strong Blinding* for evaluating the blinding

* Corresponding author. 455 E Grand Ave, Mail Stop 44-2B South San Francisco, CA 94080, USA.

property of a kit supply strategy. Sections 6 and 7, describe the Blinded Group Ordering (BGO) [7] and Extended Blinded Group Ordering (EBGO) methods and their superior blinding properties. The BGO method was originally devised to minimize drug kit wastage at low enrolling sites [7] without endangering the blind but does not appear to be widely used yet. In the *Addendum*, the BGO method for two-arm trials is proved to be strongly blinding.

For brevity, our descriptions of supply methods are limited to two-arm placebo-controlled trials in which subjects are randomized to one of the two arms in a 1:1 ratio and receive one kit per visit. However the descriptions can be generalized to more complex trials such as: multi-arm studies, studies with unequal allocation, etc.

2. Trade-off between blinding and efficiency

There is a trade-off between ensuring the blind and kit efficiency. To see this, consider the task of supplying kits in single dispensation trials, i.e., trials in which a single drug kit contains a subject's drug for the entire study. Antibiotic and vaccine trials are often single dispensation trials. Two simple supply methods, Naïve Replacement (NR) and Waste-One-Kit (WOK), illustrate the extremes of this trade-off.

- Naïve Replacement (poor blinding, high efficiency): After a kit at the site is given to a subject, this method sends a replacement kit to the site. Because the replacement kit is of the same kit type as the kit that was used, the investigator is partially unblinded when the replacement kit is used to dose a subject. That is, the investigator can conclude that the subject who is dosed with the replacement kit and the subject who was dosed with the kit that initiated the shipment of the replacement kit are on the same treatment arm.
- Waste-One-Kit (excellent blinding, low efficiency): This method protects the blind at the expense of wasting one kit per subject. It sends kits to sites in pairs: one active and one placebo. After a subject is randomized to receive either active or placebo kit, the unused kit in the pair is deactivated and a new pair of kits is sent to the site. Deactivating and not assigning the unused kit to a future subject eliminates any information regarding the used kit.

A good supply strategy will fall between these extremes and ensure the blind while minimizing the number of kits that are needed.

3. Trigger-and-resupply (TR) method

The TR method is implemented in many RTSM systems for supplying kits to sites in single dispensation studies. Each study site is assigned a *trigger level* and a *resupply level* for each kit type. When the number of kits of a certain kit type falls to or below its trigger level, a resupply shipment is sent to the site to bring the site's inventories back to the resupply level. Usually all kit types are brought up to their resupply levels, not just the kit type that has triggered the resupply. Extensions of the method exist; for example, a pre-specified number of *random kits* (randomly selected kits) can be added to a shipment in order to blind the shipment.

The trigger level is a function of accrual rate and delivery time. It is set sufficiently high so that newly randomized subjects can be dosed even while the requested stock is in transit; hence, the trigger level should be at least 1. The resupply level is set to match the accrual rate; the level is set high enough to provide drug to new subjects and to limit the frequency of resupplies but not so high to avoid risk of medication expiring and inventory exceeding site

storage capacity.

The TR method is vulnerable to partial unblinding, more so at low enrolling sites. At low enrolling sites, subject randomizations are infrequent and inventories are kept to a minimum. This provides investigators with two pieces of useful information. First, investigators can associate resupplies with the subject randomizations that have initiated the resupplies if the time between subject arrivals is longer than the time needed for resupply shipments to arrive at the sites. Second, because the inventories are kept to a minimum, the choices for parameter values are limited: usually the initial shipment will consist of 2 kits of each kit type, the trigger levels will be 1, the resupply levels will be either 2 or 3, and the number of random kits (if any) will be 1.

Here are some examples in which investigators can partially unblind kits (and hence subjects) when these two pieces of useful information are available. In the examples below, it is assumed that the initial shipment consists of 4 kits; the investigators would speculate correctly that it consists of 2 active and 2 placebo kits; and the first subject triggers a resupply.

- 1) When the first resupply shipment consists of 2 kits, the investigator would speculate correctly that the resupply level is 2 and that the first resupply shipment contains one replacement kit (a kit of the same type as the kit given to the first subject) and one random kit. Suppose the second subject receives a kit from the first resupply shipment. If another resupply is triggered, then the investigator can conclude that the second subject has received the replacement kit, both subjects are on the same treatment arm, and the random kit is for the other treatment arm. If the second subject doesn't trigger a resupply, then the remaining kit in the first resupply shipment must be of the same type as the kit used to treat the first subject.
- 2) When the first resupply shipment consists of 3 kits, the investigator would speculate correctly that the resupply level is 3. Because the site's inventory consists of 3 active and 3 placebo kits after receiving the first resupply shipment, the second subject does not trigger a resupply. If the third subject triggers a resupply, the investigator can conclude that the second and third subjects are on the same treatment arm. If the third subject does not trigger a resupply, the investigator can conclude that the second and third subjects are on opposite arms.

Blinding shipments by adding two random kits instead of one, an option offered by some RTSM systems, makes partial unblinding less likely but does not eliminate it. Furthermore adding two random kits increases site inventories and potentially the number of kits needed for the trial.

If parameter values are revealed to the investigators and the investigators can still associate resupplies with the subject randomizations that have initiated the resupplies, then partial unblinding can also occur at medium and high enrolling sites. However the risk is lower because the difference between the trigger and resupply levels has increased. If the difference between the trigger and resupply levels is maintained, then the risk remains the same.

4. Predictive Resupply (PR) method

The PR method is implemented in many RTSM systems to supply kits to sites in multiple dispensation studies. In these studies, drug kits are dispensed at more than one visit for a subject, namely at randomization and at the subject's return visits. The PR method extends the TR method with a prediction algorithm, which is used to order kits for return visits. The number of kits required for returning subjects can be predicted because subjects' visit

schedules are determined by the subjects' initial or previous visits.

Checks are made frequently, e.g. daily [6], to determine if resupply shipments are needed. The check's first step is to calculate the number of kits needed for subjects' return visits over a defined time horizon called the *check range*. A resupply is initiated if a site's inventory level falls to or below this number plus the TR trigger level. This check is made for each kit type and a resupply is initiated if the check is triggered for any kit type. The resupply shipment contains kits needed for subjects' return visits over the *prediction window* (typically a longer time horizon than the *check range*) as well as kits to meet the TR resupply levels. Like the TR method, random kits could be added to blind shipments; e.g., when the resupply shipment contains kits for a single subject.

Thus the TR method maintains a *buffer stock* of kits for subjects' first visit (all kits in the initial shipment belong to the buffer stock) and the prediction algorithm maintains a *prediction stock* of kits for subjects' return visits. However, in order to reduce inventories at sites and the number of kits needed for the trial, RTSM systems typically do not differentiate between the buffer and prediction stocks when dispensing kits to subjects. This practice and inadequacies inherited from the TR method can cause partial unblinding in the PR method, for example:

- 1) Consider a site receiving an initial shipment of 4 kits (2 active and 2 placebo kits). If a subject receives his/her first and second kit from the initial shipment, the investigator can conclude that the remaining two kits in the initial shipment are of the opposite type of the kits received by the first subject.
- 2) Consider a site with two active subjects. Suppose the site periodically receives shipments of two kits, one for each subject. If one subject leaves the trial prior to using his/her kit and both kits in the same shipment are used to dose the remaining subject, the investigator will learn that the two subjects are on the same treatment arm.
- 3) Consider a site with two active subjects. Suppose the investigator detects an increasing inventory of kits. The investigator can conclude correctly that resupply shipments contain random kits that can't be used to dose any of the subjects and that the two subjects are on the same treatment arm.
- 4) If parameter values are revealed to the investigators, the PR method could be susceptible to partial unblinding in the same way as the TR method in single dispensation trials. Hence, like the TR method, a site's parameters should be kept secret from the site.

5. Strong blinding, a criterion for evaluating strength of blinding in a clinical trial

In this section, we introduce *strong blinding* as a measure of the amount of information a supply method is able to shield from investigators in a blinded trial.

Definition 5.1: A supply method is said to be *strongly blinding at level n* , $n \geq 1$, if after being fully unblinded to any $n - 1$ kits at the site, an investigator is unable to:

- a. Fully unblind a kit, or
- b. Partially unblind two kits,

among the remaining kits (i.e., other than the $n - 1$ kits that are already fully unblinded at the site).

A supply method is strongly blinding at level 1 if it does not allow full unblinding and partial unblinding, and is strongly blinding at level 2 if it does not allow further full unblinding and partial unblinding even after the investigator is fully unblinded to

one kit. Higher levels of strong blinding impose more stringent blinding than lower levels.

The NR, TR, and PR methods are not strongly blinding at any level because partial unblinding can occur without the investigator knowing the kit types of any kits. The WOK method is strongly blinding at level infinity.

The criterion strong blinding can in fact be applied to subject randomization as well by replacing drug kits with subjects.

Definition 5.2: A randomization method is said to be *strongly blinding at level n* , $n \geq 1$, if after being fully unblinded to any $n - 1$ subjects at the site, an investigator is unable to:

- c. Fully unblind a subject, or
- d. Partially unblind two subjects,

among the remaining subjects (i.e., other than the $n - 1$ subjects that are already fully unblinded at the site).

In studies where subjects are randomized in a 1:1 ratio to receive one of the two treatment options, permuted block randomization with a block size of 4 and Big Stick randomization [8] with barrier $a = 2$ are strongly blinding at level 2, and complete randomization and Biased Coin randomization [9] are strongly blinding at level infinity. The level of strong blinding is of interest in blinded clinical trials because it is not uncommon for an investigator to be unblinded to a subject's treatment assignment for safety reasons, or for an investigator to learn a subject's treatment assignment from the subject's adverse event or efficacy profile.

6. Blinded Group Ordering (BGO)

The BGO method is a supply method for single dispensation trials. The idea behind this method is to use a random kit to blind a site's inventory instead of its shipments in order to minimize site inventories and the total number of kits required for the trial. The method has two parameters: the *supply level k* and the *supply period j* . The supply level determines the inventory level at a site after receiving the initial supply or a resupply shipment and the supply period determines how often resupply shipments are sent to the site.

For a two-arm trial with equal allocation, the method starts with an initial shipment of $2k + 1$ kits: k active kits, k placebo kits and 1 random kit. Resupply shipments, each consisting of j kits, are sent after dosing every j subjects to bring the site's inventory back to $2k + 1$ kits. Resupply shipments are configured so that after a resupply, the site again has either k active kits and $k + 1$ placebo kits, or $k + 1$ active kits and k placebo kits. To avoid stockout of any kit type at the site, we only consider $0 < j < k$.

Depending on the inventory at the site, the configuration of the resupply shipment may be deterministic or stochastic. For example, consider $k = 2$ and $j = 1$. Suppose the initial shipment consists of 2 active kits and 3 placebo kits. If the first subject at the site is assigned to the active arm, the resupply shipment has to be an active kit; on the other hand, if the subject is assigned to the placebo arm, then the resupply shipment can be either an active kit or a placebo kit, e.g. each with probability 1/2.

It should be noted that to adapt to changes in the site's accrual rate during the trial, both supply level k and supply period j can be adjusted. To increase the supply level from k to $k + 1$, simply add 1 active and 1 placebo to the resupply shipment. To decrease k to $k - 1$, send a resupply shipment of $j - 1$ kits so that after the shipment arrives there are at least $k - 1$ kits of each type at the site, and then deactivate a kit so that the inventory at the site consists of $2k - 1$ kits with at least $k - 1$ kits of each kit type. To adjust j , the change can be made after the site's inventory is replenished and has $2k + 1$ kits.

Furthermore, the BGO method can be modified to reduce the number of kits needed at study startup. The modified strategy sends $2k$ kits (k active and k placebo kits) instead of $2k + 1$ kits in the initial shipment and sends $j + 1$ kits, instead of j kits, in the first resupply shipment. After receiving the first resupply shipment, the site will have $2k + 1$ kits: k kits of one type and $k + 1$ kits of the other, and the usual strategy begins. Another advantage of this modification is that if no subjects are enrolled at a site, $2k$ kits are wasted instead of $2k + 1$.

In the [Addendum](#) we prove that for two-arm single dispensation trials, the BGO method with $j = 1$ is strongly blinding at level 3 and the BGO method with $j > 1$ is strongly blinding at level $j + 1$.

7. Extended Blinded Group Ordering (EBGO)

In this section we introduce the EBGO method, which extends the BGO method to multiple dispensation studies. The EBGO uses the BGO method to maintain a buffer stock of kits for subjects' first visit and the prediction algorithm (from the PR method) to maintain a prediction stock of kits for subjects' subsequent visits. Unlike how the PR method is commonly implemented, the buffer stock and the prediction stock are not mixed together. That is, kits in the buffer stock are only used to dose newly randomized subjects and kits in the prediction stock are only used to dose returning subjects. Furthermore each kit in the prediction stock is pre-assigned to a subject when it is shipped to the site; if the subject is discontinued from treatment before the kit is used, then the kit is deactivated.

This distinction between the buffer stock and the prediction stock and the pre-assignments of kits in the prediction stock to subjects are made in the RTSM system, but should not be visible to the sites until the kits are given to subjects for dosing.

The EBGO method is strongly blinding for subjects (as defined in Definition 5.2) at the same level as the underlying BGO method is strongly blinding for kits (as defined in Definition 5.1). For example if the underlying BGO method is strongly blinding at level 3 for kits, then the EBGO method is also strongly blinding at level 3, but for subjects.

8. Discussion

Our investigation of supply strategies started when a statistician asked us how to select TR parameters that would ensure the blind. It was clear that adding more random kits to smaller shipments increased the difficulty in detecting partial unblinding. However to our surprise we were not able to find parameter settings that we could demonstrate did not permit partial unblinding. This led us to investigate the blinding properties of alternative supply methods, including the BGO and EBGO methods, and develop the strong blinding criterion.

We show that for two-arm trials the BGO and EBGO methods are strongly blinding at a level of at least 3 even when the site personnel and sponsor's clinical trial managers are aware of the methods' parameters. Because the TR and PR methods permit partial unblinding, we conclude that the blinding properties of the BGO and EBGO methods are superior to those of the TR and PR methods.

Both blinding and efficiency are important properties of a supply method. Efficiency includes keeping to a minimum the total number of kits required for the study, the number or frequency of resupply shipments, the number of initial kits to start the study and the minimization of site inventories. Transitioning from the widely used TR and PR methods to the BGO and EBGO methods will require that the efficiency of the BGO and EBGO methods be similar to or better than the TR and PR methods. We believe that this may indeed often be the case because the TR and PR methods add

random kits to blind small shipments or shipments that may contain kits for a single subject while the BGO and EBGO methods add random kits to the initial shipments to blind sites' inventories. Therefore when the number of small shipments or shipments that may contain kits for a single subject exceeds the number of sites, the efficiency of the BGO and EBGO methods should be as good as or better than the TR and PR methods. However to thoroughly compare efficiencies of different supply methods for a given study design, simulations should be performed.

When a site has a small number of subjects who are still on treatment and if all of them are on the same treatment arm, it becomes difficult to blind the investigator by blinding shipments using random kits. This is due to the scenario as described in the third example in Section 4, which could potentially occur. An increasing inventory of kits at the site could lead the investigator to conclude correctly that all the active subjects are on the same treatment arm and the resupply shipments contain random kits that can't be used to dose any of these subjects. Therefore instead of blinding the shipments, blinding the site's inventory will be a better alternative to protect the blind, which is the approach taken by the BGO and EBGO methods.

Another advantage of the BGO and EBGO methods is that because there is no buildup of inventory at the site due to added random kits, the size of inventory at the site is controlled. By keeping more kits (kits that would otherwise be sent to the site as random kits) at the depot rather than at the site, the BGO and EBGO methods improve the flexibility of supply chain.

As described earlier, the EBGO method pre-assigns kits in the prediction stock to subjects in order to avoid unblinding. This pre-assignment could result in drug kit wastage up to the amount of supply need over the prediction window per subject, especially for trials in which the treatment duration is not fixed – for example, subjects are often treated until disease progression in oncology studies. The cost of pre-assignment, which fortunately does not need to be paid at the beginning of a trial, has to be weighed against the risk of unblinding.

We hope that the superior blinding properties of the BGO and EBGO methods as demonstrated in this paper coupled with a confirmation of similar or improved efficiency, may convince company sponsors and RTSM system providers to transition to the BGO and EBGO methods to supply kits for suitable trials.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Addendum

In this section we show that the BGO method for two-arm single dispensation trials (with $k > j > 0$) is strongly blinding at a level that is at least 3.

Theorem A1

- The BGO method with $j = 1$ ($k > j$) is strongly blinding at level 3 but not at level 4.
- The BGO method with $j > 1$ ($k > j$) is strongly blinding at level $j + 1$ but not at level $j + 2$.

(Note that the BGO method with $j = 1$ is strongly blinding at level 3, same as the BGO method with $j = 2$.)

Proof of Theorem A1

The proof is based on the construction of “solutions” and to

show that this solution space is “large”. Each solution is an assignment of kits to kit types that are consistent with all the information known to the investigator. Specifically to show that the BGO method is strongly blinding at level n , we show that for any $n + 1$ kits at the site, after unblinding the investigator to the first $n - 1$ kits, there exist, regardless of the kit types of the first $n - 1$ kits, the following four solutions:

- S1 A solution in which kit n is assigned to active,
- S2 A solution in which kit n is assigned to placebo,
- S3 A solution in which kits n and $n + 1$ have the same kit type, and
- S4 A solution in which kits n and $n + 1$ have opposite kit types.

When all four solutions exist, it is not possible for the investigator to unblind kit n or to partially unblind kits n and $n + 1$ even after learning the kit types of the first $n - 1$ kits. We will use the notation X_1, \dots, X_{n-1} for the first $n - 1$ kits, and X_n and X_{n+1} for kit n and kit $n + 1$ respectively.

BGO method with $j = 1$

After the initial shipment and every resupply shipment, the site's inventory will consist of $2k + 1$ kits: k active kits, k placebo kits and one random kit. Resupply shipments, each consisting of a single kit, are sent after dosing each subject.

During the trial, information in the site's inventory available to the investigator can be summarized in an inventory table, where columns represent the “snapshots” of the inventory after each supply (initial supply or a resupply). Unused kits from the previous column remain in the same rows.

Consider the following hypothetical example (using BGO method with $k = 2$ and $j = 1$): a site enrolled a total of 10 subjects during the course of the study (thus a total of 15 kits have been sent to the site) and kits #3, 6, 7, 2, 1, 5, 9, 4, 12, and 11 were given in this order to dose the 10 subjects. The inventory table as described above is shown in Table A1.

Table A1

Inventory table, showing the kit numbers of kits in the inventory, for the hypothetical example. Kits that are new to the inventory are in bold.

Initial inventory	Inventory after i th resupply										Final inventory
	$i = 1$	$i = 2$	$i = 3$	$i = 4$	$i = 5$	$i = 6$	$i = 7$	$i = 8$	$i = 9$	$i = 10$	
1	1	1	1	1	10	10	10	10	10	10	10
2	2	2	2	9	9	9	12	12	14	14	14
3	6	7	8	8	8	8	8	8	8	8	8
4	4	4	4	4	4	4	4	13	13	13	13
5	5	5	5	5	5	11	11	11	11	15	15

The only constraint on the inventory imposed by the BGO method is that each column of Table A1 must contain a minimum of 2 placebo and 2 active kits. There are many solutions that assign kits to kit types (active or placebo) that satisfy this constraint. To prove Theorem A1 for $j = 1$, we will need two classes of solutions, “trivial” solutions and “change over” solutions.

A trivial solution consists of k fixed active rows, k fixed placebo rows and one free row. All kits in a fixed row are assigned to the same kit type, either active or placebo. Kits in the free row can be assigned to either kit type. For example a trivial solution for the hypothetical example sets rows 1 and 2 to active, rows 3 and 4 to placebo, and row 5 as a free row. That is, kits (1, 2, 9, 10, 12, 14) are active kits, kits (4, 6, 7, 8, 13) are placebo kits, and kits (5, 11, 15) in row 5 can be any of the 8 combinations of kit types, i.e., (a, a, a) , (a, a, p) , ..., (p, p, p) , where a = active and p = placebo.

A change over solution consists of $2k - 1$ fixed rows and two “change” rows: an “active” change row (which is a change row that begins with an active kit) and a “placebo” change row (which is a change row that begins with a placebo kit). A change row has a single change point in the row. The row begins with one type of kit and then changes to the other type. For example in the hypothetical example, if kits 2 and 9 are set to active and kits 12 and 14 are set to placebo in row 2, then row 2 is an active change row with the change point in between kits 9 and 12. The key observation is that the two change rows will contribute one active and one placebo kit to the inventory before the first change point and after the second change point. Between the two change points, the two change rows will either contribute two active kits or two placebo kits depending on which change point occurs first. If the change rows contribute two active (placebo) kits between the two change points, then setting k fixed rows to placebo (active) and $k - 1$ fixed row to active (placebo) will satisfy the constraints.

To prove that BGO with $j = 1$ is strongly blinding at level 3, we need to obtain solutions S1–4 for an arbitrary selection of 4 kits (X_1, X_2, X_3, X_4) from the site, regardless of the kit types given for X_1 and X_2 .

Finding solutions S1 and S2 is equivalent to demonstrating that it is possible to assign kits (X_1, X_2, X_3) to any of the 8 combinations of kit types regardless of the locations of the three kits (in the inventory table). Only trivial solutions are needed to obtain solutions S1 and S2. There are 3 cases:

1. The three kits are in distinct rows – each of the three rows can be either a fixed active row or a fixed placebo row.
2. The three kits are in the same row – make this row the free row to obtain all eight treatment combinations.
3. Two kits are in one row and the other kit is in another – make the row with the two kits the free row.

To obtain solutions S3 and S4, there are 6 cases to consider:

1. The four kits are in the same row.

2. The four kits are in distinct rows.
3. Three kits are in one row and the other kit is in another row.
4. Two kits are in one row and the other kits are in distinct rows.
5. X_1 and X_3 are in one row and X_2 and X_4 are in another row.
6. X_1 and X_2 are in one row and X_3 and X_4 are in another row.

It should be noted that unlike finding S1 and S2, finding solutions S3 and S4 is not the same as demonstrating the possibility of assigning kits (X_1, X_2, X_3, X_4) to any of the 16 combinations of kit types regardless of the locations of the four kits. For example, suppose kits (X_1, X_2, X_3, X_4) are four of the initial kits, then it is not possible to assign all of them to active for the BGO method with $k = 2$.

Without going into details, we claim that trivial solutions can be easily obtained except for case 6 where difficulty arises in

establishing a trivial solution in which X_1 and X_2 are of opposite kit types and kits X_3 and X_4 are of opposite kit types as there is only one free row. Here a “change over” solution is needed.

To construct the change over solution for case 6, make the two rows containing kits (X_1, X_2, X_3, X_4) two change rows and for each change row, set the change point to any place between the two selected kits in the row. Without loss of generality, assume that X_2 arrives at the site after X_1 and X_4 arrives at the site after X_3 . If X_1 is an active (placebo) kit, make the row containing X_3 and X_4 a placebo (active) change row. The selected kits in each of the change rows will have opposite kit types (because they fall on opposite sides of the change point) thus completing the proof of the assertion that BGO method with $j = 1$ is strongly blinding at level 3.

The BGO method with $j = 1$ is not strongly blinding at level 4 because suppose that the first 2 kits used at the site come from the initial shipment, if the investigator learns that both are active kits and the first resupply is a placebo kit, then the investigator can conclude with certainty that the second resupply is an active kit.

BGO method with $j > 1$

After the initial shipment and every resupply shipment, a site's inventory will consist of $2k + 1$ kits: k active kits, k placebo kits and one random kit. Each resupply shipment consists of j kits and is sent after every j subjects. Similar to the previous case of $j = 1$, information in the site's inventory can be summarized in a table with $2k + 1$ rows. However because the size of the resupply shipments is greater than 1, the inventory table is not unique (kits sent in the same shipment can be permuted among themselves). In particular, if N resupply shipments have been sent to the site, there are $(j!)^N$ distinct inventory tables up to permutation of rows.

To prove that BGO with $j > 1$ is strongly blinding at level $j + 1$, we need to obtain solutions S1–4 for an arbitrary selection of $j + 2$ kits (X_1, \dots, X_{j+2}) from the site, regardless of the kit types given for the first j kits (X_1, \dots, X_j).

The following claim can be made: for any j kits, there exists an inventory table such that each of the j kits is in a distinct row. This is true because for each kit that arrives at the site, there are at a minimum j possible rows to assign the kit to. Therefore an inventory table in which the first j kits (X_1, \dots, X_j) are in distinct rows can always be found.

There are two places for X_{j+1} : in a new row or in a previously occupied row. Therefore there are two cases to consider for the first $j + 1$ kits (X_1, \dots, X_{j+1}):

1. All kits are in distinct rows.
2. Two kits are in one row and the other kits are in distinct rows.

Trivial solutions can be easily obtained for S1 and S2.

Placement of X_{j+1} and X_{j+2} leads to the following 4 cases (after possibly relabeling the kits):

1. All kits are in distinct rows.
2. Three kits are in one row and the other $j - 1$ kits are in distinct rows.
3. Two kits are in one row and other j kits are in distinct rows.
4. Kits X_1 and X_{j+1} are in one row, kits X_2 and X_{j+2} are in one row, and the other $j - 2$ kits are in distinct rows.

It can be seen that trivial solutions can also be obtained for S3 and S4.

The BGO method with $j > 1$ is not strongly blinding at level $j + 2$ because if the first j kits used at the site come from the initial shipment, the investigator learns that they are active kits, and a kit in the first resupply is placebo, then the investigator can conclude that the remaining kits in the first resupply must be active kits. This counterexample does not apply to the BGO method with $j = 1$ because each resupply has only one kit. No kits remain in the first resupply after the placebo kit is unblinded.

It should be noted that in the proof of Theorem A1, it is assumed that all kits of the appropriate kit type have a non-zero probability of being selected when a subject needs a kit. If this is not the case, e.g. kits in the earliest shipments are assigned first (i.e., FIFO), additional constraints on the solution space may be introduced and Theorem A1 no longer holds.

References

- [1] M. Lang, R. Wood, D. McEntegart, Double randomisation: protecting the blind, *Good Clin. Pract.* 12.11 (2005) 10–13. Nottingham: T&F Informa UK Ltd.
- [2] O.M. Kuznetsova, Why permutation is even more important in IVRS drug codes schedule generation than in patient randomization schedule generation, *Control. Clin. Trials* 22.1 (2001) 69–71.
- [3] D.A. Coleman, R. Yu, The other randomization—methods for labeling drug kits, *Contemp. Clin. Trials* 38.2 (2014) 270–274.
- [4] S. Day, Letters to the editors, *Control. Clin. Trials* 22 (2001) 264–265.
- [5] B. Byrom, Managing the medication supply chain process using interactive voice response system, *Life Sci. Today* 3 (2002) 16–19.
- [6] M. Peterson, B. Byrom, N. Dowlmanb, D. McEntegart, Optimizing clinical trial supply requirements: simulation of computer-controlled supply chain management, *Clin. Trials* 1.4 (2004) 399–412.
- [7] S. Waters, I. Dowlman, K. Drake, L. Gamble, M. Lang, D. McEntegart, Enhancing control of the medication supply chain in clinical trials managed by interactive voice response systems, *Drug Inf. J.* 44.6 (2010) 727–740.
- [8] J.F. Soares, C.F.J. Wu, Some restricted randomization rules in sequential designs, *Commun. Stat. Theory Methods* 12.17 (1983) 2017–2034.
- [9] B. Efron, Forcing a sequential experiment to be balanced, *Biometrika* 58.3 (1971) 403–417.